

RAW SEQUENCE LISTING

The Biotechnology Systems Branch of the Scientific and Technical Information Center (STIC) no errors detected.

Application Serial Number: 10/528,960
Source: Pgt/10
Date Processed by STIC: 4/1/05

ENTERED



PCT

RAW SEQUENCE LISTING

DATE: 04/01/2005

PATENT APPLICATION: US/10/528,960

TIME: 10:34:40

Input Set : A:\0185946SeqList.txt

Output Set: N:\CRF4\04012005\J528960.raw

3 <110> APPLICANT: DSM IP ASSETS B.V.
 5 <120> TITLE OF INVENTION: Enone reductase gene
 7 <130> FILE REFERENCE: NDR5231
 C--> 9 <140> CURRENT APPLICATION NUMBER: US/10/528,960
 C--> 10 <141> CURRENT FILING DATE: 2005-03-22
 12 <150> PRIOR APPLICATION NUMBER: EP 02021098.5
 13 <151> PRIOR FILING DATE: 2002-09-23
 15 <160> NUMBER OF SEQ ID NOS: 32
 17 <170> SOFTWARE: PatentIn version 3.2
 19 <210> SEQ ID NO: 1
 20 <211> LENGTH: 1212
 21 <212> TYPE: DNA
 22 <213> ORGANISM: Candida kefyr
 25 <220> FEATURE:
 26 <221> NAME/KEY: CDS
 27 <222> LOCATION: (1)..(1212)
 29 <400> SEQUENCE: 1

30	atg	tcg	tac	atg	aac	ttt	gac	cct	aag	cca	ttg	gga	gac	acc	aat	atc	48
31	Met	Ser	Tyr	Met	Asn	Phe	Asp	Pro	Lys	Pro	Leu	Gly	Asp	Thr	Asn	Ile	
32	1				5					10				15			
34	ttc	aag	cca	atc	aag	atc	ggt	aac	aat	gag	cta	aaa	cac	aga	gta	gtc	96
35	Phe	Lys	Pro	Ile	Lys	Ile	Gly	Asn	Asn	Glu	Leu	Lys	His	Arg	Val	Val	
36				20				25					30				
38	atg	cca	gca	ttg	act	aga	atg	aga	gcc	att	gca	cca	gga	aac	atc	cca	144
39	Met	Pro	Ala	Leu	Thr	Arg	Met	Arg	Ala	Ile	Ala	Pro	Gly	Asn	Ile	Pro	
40			35				40				45						
42	aac	act	gaa	tgg	gcc	gag	gaa	tac	tac	aga	caa	cgt	tct	caa	tac	cct	192
43	Asn	Thr	Glu	Trp	Ala	Glu	Glu	Tyr	Tyr	Arg	Gln	Arg	Ser	Gln	Tyr	Pro	
44		50				55				60							
46	ggt	acc	ctt	att	atc	acg	gaa	ggt	act	ttc	cct	tct	gcg	caa	tca	ggt	240
47	Gly	Thr	Leu	Ile	Ile	Thr	Glu	Gly	Thr	Phe	Pro	Ser	Ala	Gln	Ser	Gly	
48	65				70				75				80				
50	ggt	tac	cca	aat	gtg	cca	ggt	atc	tgg	tcc	aaa	gag	caa	ttg	gct	gaa	288
51	Gly	Tyr	Pro	Asn	Val	Pro	Gly	Ile	Trp	Ser	Lys	Glu	Gln	Leu	Ala	Glu	
52				85				90				95					
54	tgg	aaa	aag	atc	ttc	aat	gca	atc	cat	gag	aac	aaa	tcg	ttc	gtg	tgg	336
55	Trp	Lys	Lys	Ile	Phe	Asn	Ala	Ile	His	Glu	Asn	Lys	Ser	Phe	Val	Trp	
56			100				105					110					
58	gtg	caa	ttg	tgg	gtt	cta	ggt	aga	caa	gca	tgg	cca	gaa	gtg	ttg	aag	384
59	Val	Gln	Leu	Trp	Val	Leu	Gly	Arg	Gln	Ala	Trp	Pro	Glu	Val	Leu	Lys	
60			115				120					125					
62	aag	gaa	ggt	ttg	cgt	tac	gat	agt	gct	acc	gat	gac	ttg	tac	atg	ggt	432
63	Lys	Glu	Gly	Leu	Arg	Tyr	Asp	Ser	Ala	Thr	Asp	Asp	Leu	Tyr	Met	Gly	

p.6

RAW SEQUENCE LISTING

DATE: 04/01/2005

PATENT APPLICATION: US/10/528,960

TIME: 10:34:40

Input Set : A:\0185946SeqList.txt

Output Set: N:\CRF4\04012005\J528960.raw

64	130	135	140	
66	gaa gaa gaa aaa gag cgt gcc tta aag gct aac aac cca cag cac ggt	480		
67	Glu Glu Glu Lys Glu Arg Ala Leu Lys Ala Asn Asn Pro Gln His Gly			
68	145 150 155 160			
70	atc acc aag gaa gaa atc aag cag tac atc aag gag tac gtg gat gct	528		
71	Ile Thr Lys Glu Glu Ile Lys Gln Tyr Ile Lys Glu Tyr Val Asp Ala			
72	165 170 175			
74	gcc aag aaa gcc atc gat gca ggt gca gac ggt gtg caa atc cat tct	576		
75	Ala Lys Lys Ala Ile Asp Ala Gly Ala Asp Gly Val Gln Ile His Ser			
76	180 185 190			
78	gcc aac ggt tac ttg ttg aac cag ttt ttg gac cct att tct aac aac	624		
79	Ala Asn Gly Tyr Leu Leu Asn Gln Phe Leu Asp Pro Ile Ser Asn Asn			
80	195 200 205			
82	aga acc gac gag tac ggt gga tcg atc gag aac cgt gcg aga ttc act	672		
83	Arg Thr Asp Glu Tyr Gly Gly Ser Ile Glu Asn Arg Ala Arg Phe Thr			
84	210 215 220			
86	ttg gaa gtg gtc gat gcc gtt gtc gat gca gtt ggt gcc gaa aga acc	720		
87	Leu Glu Val Val Asp Ala Val Val Asp Ala Val Gly Ala Glu Arg Thr			
88	225 230 235 240			
90	tcc atc aga ttc tct cca tac ggt act ttt ggt acc atg tcc ggt ggt	768		
91	Ser Ile Arg Phe Ser Pro Tyr Gly Thr Phe Gly Thr Met Ser Gly Gly			
92	245 250 255			
94	gag aac cct ggc atc gtt gct caa tat gca tac gtc att ggt gag ttg	816		
95	Glu Asn Pro Gly Ile Val Ala Gln Tyr Ala Tyr Val Ile Gly Glu Leu			
96	260 265 270			
98	gaa aag aga gct aga gct ggc aag aga ttg gcg ttc atc gat ttg gtc	864		
99	Glu Lys Arg Ala Arg Ala Gly Lys Arg Leu Ala Phe Ile Asp Leu Val			
100	275 280 285			
102	gag cct cgt gtg acc gac cca ttc cta cca gaa ttc gag aag tgg ttc	912		
103	Glu Pro Arg Val Thr Asp Pro Phe Leu Pro Glu Phe Glu Lys Trp Phe			
104	290 295 300			
106	aag gaa ggt acc aac gaa ttc atc tac tct atc tgg aag ggt cca gtt	960		
107	Lys Glu Gly Thr Asn Glu Phe Ile Tyr Ser Ile Trp Lys Gly Pro Val			
108	305 310 315 320			
110	ctc aga gtt ggt aac tat gct ttg gac cca gat caa gcc act ctc gac	1008		
111	Leu Arg Val Gly Asn Tyr Ala Leu Asp Pro Asp Gln Ala Thr Leu Asp			
112	325 330 335			
114	tct aag aag cct aac act ttg atc ggt tac ggt aga tcc ttc atc gcc	1056		
115	Ser Lys Lys Pro Asn Thr Leu Ile Gly Tyr Gly Arg Ser Phe Ile Ala			
116	340 345 350			
118	aac cca gac ttg gtg tac cgt ttg gaa aag ggt ttg cca ttg aac aag	1104		
119	Asn Pro Asp Leu Val Tyr Arg Leu Glu Lys Gly Leu Pro Leu Asn Lys			
120	355 360 365			
122	tat gat aga aac acc ttt tac aca ttc act aag gaa ggt tac acc gat	1152		
123	Tyr Asp Arg Asn Thr Phe Tyr Thr Phe Thr Lys Glu Gly Tyr Thr Asp			
124	370 375 380			
126	tac cca agc tac gaa gaa tcc gtc gca aag ggt tac aag aaa gag gaa	1200		
127	Tyr Pro Ser Tyr Glu Glu Ser Val Ala Lys Gly Tyr Lys Lys Glu Glu			
128	385 390 395 400			

RAW SEQUENCE LISTING

DATE: 04/01/2005

PATENT APPLICATION: US/10/528,960

TIME: 10:34:40

Input Set : A:\0185946SeqList.txt

Output Set: N:\CRF4\04012005\J528960.raw

1212

```

130 aag aag tac taa
131 Lys Lys Tyr
135 <210> SEQ ID NO: 2
136 <211> LENGTH: 403
137 <212> TYPE: PRT
138 <213> ORGANISM: Candida kefyr
140 <400> SEQUENCE: 2
142 Met Ser Tyr Met Asn Phe Asp Pro Lys Pro Leu Gly Asp Thr Asn Ile
143 1 5 10 15
146 Phe Lys Pro Ile Lys Ile Gly Asn Asn Glu Leu Lys His Arg Val Val
147 20 25 30
150 Met Pro Ala Leu Thr Arg Met Arg Ala Ile Ala Pro Gly Asn Ile Pro
151 35 40 45
154 Asn Thr Glu Trp Ala Glu Glu Tyr Tyr Arg Gln Arg Ser Gln Tyr Pro
155 50 55 60
158 Gly Thr Leu Ile Ile Thr Glu Gly Thr Phe Pro Ser Ala Gln Ser Gly
159 65 70 75 80
162 Gly Tyr Pro Asn Val Pro Gly Ile Trp Ser Lys Glu Gln Leu Ala Glu
163 85 90 95
166 Trp Lys Lys Ile Phe Asn Ala Ile His Glu Asn Lys Ser Phe Val Trp
167 100 105 110
170 Val Gln Leu Trp Val Leu Gly Arg Gln Ala Trp Pro Glu Val Leu Lys
171 115 120 125
174 Lys Glu Gly Leu Arg Tyr Asp Ser Ala Thr Asp Asp Leu Tyr Met Gly
175 130 135 140
178 Glu Glu Glu Lys Glu Arg Ala Leu Lys Ala Asn Asn Pro Gln His Gly
179 145 150 155 160
182 Ile Thr Lys Glu Glu Ile Lys Gln Tyr Ile Lys Glu Tyr Val Asp Ala
183 165 170 175
186 Ala Lys Lys Ala Ile Asp Ala Gly Ala Asp Gly Val Gln Ile His Ser
187 180 185 190
190 Ala Asn Gly Tyr Leu Leu Asn Gln Phe Leu Asp Pro Ile Ser Asn Asn
191 195 200 205
194 Arg Thr Asp Glu Tyr Gly Gly Ser Ile Glu Asn Arg Ala Arg Phe Thr
195 210 215 220
198 Leu Glu Val Val Asp Ala Val Val Asp Ala Val Gly Ala Glu Arg Thr
199 225 230 235 240
202 Ser Ile Arg Phe Ser Pro Tyr Gly Thr Phe Gly Thr Met Ser Gly Gly
203 245 250 255
206 Glu Asn Pro Gly Ile Val Ala Gln Tyr Ala Tyr Val Ile Gly Glu Leu
207 260 265 270
210 Glu Lys Arg Ala Arg Ala Gly Lys Arg Leu Ala Phe Ile Asp Leu Val
211 275 280 285
214 Glu Pro Arg Val Thr Asp Pro Phe Leu Pro Glu Phe Glu Lys Trp Phe
215 290 295 300
218 Lys Glu Gly Thr Asn Glu Phe Ile Tyr Ser Ile Trp Lys Gly Pro Val
219 305 310 315 320
222 Leu Arg Val Gly Asn Tyr Ala Leu Asp Pro Asp Gln Ala Thr Leu Asp
223 325 330 335

```

RAW SEQUENCE LISTING

DATE: 04/01/2005

PATENT APPLICATION: US/10/528,960

TIME: 10:34:40

Input Set : A:\0185946SeqList.txt

Output Set: N:\CRF4\04012005\J528960.raw

```

226 Ser Lys Lys Pro Asn Thr Leu Ile Gly Tyr Gly Arg Ser Phe Ile Ala
227           340           345           350
230 Asn Pro Asp Leu Val Tyr Arg Leu Glu Lys Gly Leu Pro Leu Asn Lys
231           355           360           365
234 Tyr Asp Arg Asn Thr Phe Tyr Thr Phe Thr Lys Glu Gly Tyr Thr Asp
235           370           375           380
238 Tyr Pro Ser Tyr Glu Glu Ser Val Ala Lys Gly Tyr Lys Lys Glu Glu
239 385           390           395           400
242 Lys Lys Tyr
246 <210> SEQ ID NO: 3
247 <211> LENGTH: 12
248 <212> TYPE: PRT
249 <213> ORGANISM: Candida kefyr
251 <400> SEQUENCE: 3
253 Pro Leu Gly Asp Thr Asn Ile Phe Lys Pro Ile Lys
254 1           5           10
257 <210> SEQ ID NO: 4
258 <211> LENGTH: 50
259 <212> TYPE: PRT
260 <213> ORGANISM: Candida kefyr
262 <400> SEQUENCE: 4
264 His Arg Val Val Met Pro Ala Leu Thr Arg Met Arg Ala Ile Ala Pro
265 1           5           10           15
268 Gly Asn Ile Pro Asn Thr Glu Trp Ala Glu Glu Tyr Tyr Arg Gln Arg
269           20           25           30
272 Ser Gln Tyr Pro Gly Thr Leu Ile Ile Thr Glu Gly Thr Phe Pro Ser
273           35           40           45
276 Val Gln
277           50
280 <210> SEQ ID NO: 5
281 <211> LENGTH: 8
282 <212> TYPE: PRT
283 <213> ORGANISM: Candida kefyr
285 <400> SEQUENCE: 5
287 Glu Gln Leu Ala Glu Trp Lys Lys
288 1           5
291 <210> SEQ ID NO: 6
292 <211> LENGTH: 9
293 <212> TYPE: PRT
294 <213> ORGANISM: Candida kefyr
296 <400> SEQUENCE: 6
298 Ile Phe Asn Ala Ile His Glu Asn Lys
299 1           5
302 <210> SEQ ID NO: 7
303 <211> LENGTH: 18
304 <212> TYPE: PRT
305 <213> ORGANISM: Candida kefyr
308 <220> FEATURE:
309 <221> NAME/KEY: MISC_FEATURE

```

RAW SEQUENCE LISTING

DATE: 04/01/2005

PATENT APPLICATION: US/10/528,960

TIME: 10:34:40

Input Set : A:\0185946SeqList.txt

Output Set: N:\CRF4\04012005\J528960.raw

```

310 <222> LOCATION: (15)..(15)
312 <220> FEATURE:
313 <221> NAME/KEY: MISC_FEATURE
314 <222> LOCATION: (15)..(15)
315 <223> OTHER INFORMATION: Xaa is unknown
317 <400> SEQUENCE: 7
W--> 319 Ser Phe Val Trp Val Gln Leu Trp Val Leu Gly Arg Gln Ala Xaa Pro
320 1 5 10 15
323 Glu Val
327 <210> SEQ ID NO: 8
328 <211> LENGTH: 19
329 <212> TYPE: PRT
330 <213> ORGANISM: Candida kefyr
332 <400> SEQUENCE: 8
334 Glu Gly Leu Arg Tyr Asp Ser Ala Phe Asp Asp Leu Tyr Met Gly Glu
335 1 5 10 15
338 Glu Glu Lys
342 <210> SEQ ID NO: 9
343 <211> LENGTH: 10
344 <212> TYPE: PRT
345 <213> ORGANISM: Candida kefyr
347 <400> SEQUENCE: 9
349 Ala Asn Asn Pro Gln His Gly Ile Thr Lys
350 1 5 10
353 <210> SEQ ID NO: 10
354 <211> LENGTH: 7
355 <212> TYPE: PRT
356 <213> ORGANISM: Candida kefyr
358 <400> SEQUENCE: 10
360 Glu Tyr Val Asp Ala Ala Lys
361 1 5
364 <210> SEQ ID NO: 11
365 <211> LENGTH: 50
366 <212> TYPE: PRT
367 <213> ORGANISM: Candida kefyr
370 <220> FEATURE:
371 <221> NAME/KEY: MISC_FEATURE
372 <222> LOCATION: (43)..(43)
373 <223> OTHER INFORMATION: Xaa is unknown
375 <220> FEATURE:
376 <221> NAME/KEY: MISC_FEATURE
377 <222> LOCATION: (45)..(45)
378 <223> OTHER INFORMATION: Xaa is unknown
380 <220> FEATURE:
381 <221> NAME/KEY: MISC_FEATURE
382 <222> LOCATION: (47)..(48)
383 <223> OTHER INFORMATION: Xaa is unknown
385 <400> SEQUENCE: 11
387 Ala Ile Asp Ala Gly Ala Asp Gly Val Gln Ile His Ser Ala Asn Gly

```

RAW SEQUENCE LISTING ERROR SUMMARY
PATENT APPLICATION: US/10/528,960

DATE: 04/01/2005
TIME: 10:34:41

Input Set : A:\0185946SeqList.txt
Output Set: N:\CRF4\04012005\J528960.raw

Please Note:

Use of n and/or Xaa have been detected in the Sequence Listing. Please review the Sequence Listing to ensure that a corresponding explanation is presented in the <220> to <223> fields of each sequence which presents at least one n or Xaa.

Seq#:7; Xaa Pos. 15

Seq#:11; Xaa Pos. 43, 45, 47, 48

Seq#:31; N Pos. 24

Seq#:32; N Pos. 9

VERIFICATION SUMMARY

DATE: 04/01/2005

PATENT APPLICATION: US/10/528,960

TIME: 10:34:41

Input Set : A:\0185946SeqList.txt

Output Set: N:\CRF4\04012005\J528960.raw

L:9 M:270 C: Current Application Number differs, Replaced Current Application Number

L:10 M:271 C: Current Filing Date differs, Replaced Current Filing Date

L:319 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:7 after pos.:0

L:395 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:11 after pos.:32

L:930 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:31 after pos.:0

L:958 M:341 W: (46) "n" or "Xaa" used, for SEQ ID#:32 after pos.:0